

UC Berkeley

Proposals from the Script Encoding Initiative

Title

Summary proposal to encode characters for Vedic in the BMP of the UCS

Permalink

<https://escholarship.org/uc/item/1bk045t1>

Authors

Everson, Michael
Scharf, Peter

Publication Date

2008-03-06

Peer reviewed

Universal Multiple-Octet Coded Character Set
International Organization for Standardization
Organisation internationale de normalisation
Международная организация по стандартизации

Doc Type: Working Group Document

Title: Summary proposal to encode characters for Vedic in the BMP of the UCS
Source: Ireland and UC Berkeley Script Encoding Initiative (Universal Scripts Project)
Authors: Michael Everson and Peter Scharf
Status: National Body and Liaison Contribution
Action: For consideration by JTC1/SC2/WG2 and UTC
Replaces: N3366
Date: 2008-03-06

1. Introduction. This document reflects the consensus achieved after intense discussion of the Vedic characters proposed by CDAC in India on the one hand by Everson, Scharf, et al. on the other, in a large number of documents presented since 2006. The repertoire given here comprises a set of characters which all parties agree should be encoded to enable India and Western Vedic specialists to represent the texts and critical apparatus of this important field of study.

A few outstanding issues remain but it is agreed that these might be left for further study. The characters presented in this document are, in our opinion, well-documented and appropriate for encoding.

It is worthwhile reviewing some of the issues which were previously controversial.

2. PRISHTHAMATRA E. It is our view that the best way to handle this is to encode a single vowel sign for *e* that can be used in combination with other characters for *o*, *ai*, and *au*, as presented in N3235R. We do not believe that the spoofing danger of this character is as great as has been suggested: one could recommend to registrars, for instance, that the character can be restricted from use in IDN. We propose to encode DEVANAGARI VOWEL SIGN PRISHTHAMATRA E at U+094E. Our view is that using U+0947 for this with a radical glyph difference for prishthamatra orthography would simply be shifting the spoofing concern to a far more dangerous place. Protocols could restrict the use of U+094E DEVANAGARI VOWEL SIGN PRISHTHAMATRA E; they can't do the same for U+0947 DEVANAGARI VOWEL SIGN E.

3. Spacing candrabindus and anusvaras. The characters proposed at U+1CE8..U+1CEB, U+1CED..U+1CEF, and U+A8F3..U+A8F9, were proposed as combining characters by CDAC because there should not be a linebreak before them. Because they themselves may bear diacritical marks (for example: त्रुं) and indeed because they are sometime separated (in manuscripts) by a not inconsiderable space from the characters on either side of them, they are proposed here as ordinary spacing characters. Linebreaks should not be allowed before them, however.

4. Cantillation marks for Samavedic. Ruby is a system of annotation. The combining digits and letters used in Sāmavedic are a system of cantillation marks, just like those used in Hebrew. The fact that the glyphs for these marks in the Sāmavedic tradition were derived from digits and letters is incidental. They could equally well be stars and snakes and dots and double dots and tick marks. There is no “two-ness” inherent in COMBINING DEVANAGARI DIGIT TWO. It is used to indicate a *svarita* syllable, or an *udātta* syllable when followed by an *anudātta* syllable. Similarly, there is no “ka-ness” inherent in COMBINING

DEVANAGARI DIGIT KA. It indicates *anudātta* preceding an independent *svarita*. Further, there are non-numeric and non-literal tone marks also used in Sāmaveda—KARSHANA, SHARA, PRENKHA—and it does not make sense to have some cantillation marks represented as combining characters and others as Ruby annotation. This is a *Gedankenexperiment* gone overboard. In point of fact, we can easily imagine text-to-speech software that could sing Sāmavedic text marked up with its in-band diacritics. Out-of-band Ruby annotations are ancillary footnotes to text—the Sāmavedic cantillation marks are essential diacritics, part of the text, not annotations to it.

Much ink has been spilt over this, but it has been based on a superficial analysis of the glyphs. What is more puzzling is that Latin and Cyrillic combining letters have been encoded, without the same level of push-back. We believe that we must reject Ruby as a means for representing Sāmavedic cantillation marks. Ruby is appropriate and necessary for Jaiminīya. It is not for Sāmaveda.

5. Unicode Character Properties. Character properties are proposed here.

```
0900;DEVANAGARI SIGN INVERTED CANDRABINDU;Mn;0;NSM;;;;;N;;;;;
094E;DEVANAGARI VOWEL SIGN PRISHTHAMATRA E;Mc;0;L;;;;;N;;;;;
0955;DEVANAGARI VOWEL SIGN CANDRA LONG E;Mn;0;NSM;;;;;N;;;;;
0973;DEVANAGARI SIGN PUSHPIKA;Po;0;L;;;;;N;;;;;
0974;DEVANAGARI CARET;Po;0;L;;;;;N;;;;;
0979;DEVANAGARI LETTER ZHA;Lo;0;L;;;;;N;;;;;
097A;DEVANAGARI LETTER HEAVY YA;Lo;0;L;;;;;N;;;;;
1CD0;VEDIC TONE KARSHANA;Mn;230;NSM;;;;;N;;;;;
1CD1;VEDIC TONE SHARA;Mn;230;NSM;;;;;N;;;;;
1CD2;VEDIC TONE PRENKHA;Mn;230;NSM;;;;;N;;;;;
1CD3;VEDIC TONE NIHSHVASA;Po;0;L;;;;;N;;;;;
1CD5;VEDIC TONE YAJURVEDIC AGGRAVATED INDEPENDENT SVARITA;Mn;220;NSM;;;;;N;;;;;
1CD6;VEDIC TONE YAJURVEDIC INDEPENDENT SVARITA;Mn;220;NSM;;;;;N;;;;;
1CD7;VEDIC TONE YAJURVEDIC KATHAKA INDEPENDENT SVARITA;Mn;220;NSM;;;;;N;;;;;
1CD8;VEDIC TONE CANDRA BELOW;Mn;220;NSM;;;;;N;;;;;
1CD9;VEDIC TONE YAJURVEDIC KATHAKA INDEPENDENT SVARITA SCHROEDER;Mn;220;NSM;;;;;N;;;;;
1CDA;VEDIC TONE DOUBLE SVARITA;Mn;230;NSM;;;;;N;;;;;
1CDB;VEDIC TONE TRIPLE SVARITA;Mn;230;NSM;;;;;N;;;;;
1CDC;VEDIC TONE KATHAKA ANUDATTA;Mn;220;NSM;;;;;N;;;;;
1CDD;VEDIC TONE DOT BELOW;Mn;220;NSM;;;;;N;;;;;
1CDE;VEDIC TONE TWO DOTS BELOW;Mn;220;NSM;;;;;N;;;;;
1CDF;VEDIC TONE THREE DOTS BELOW;Mn;220;NSM;;;;;N;;;;;
1CF0;VEDIC TONE RIGVEDIC KASHMIRI INDEPENDENT SVARITA;Mn;230;NSM;;;;;N;;;;;
1CF1;VEDIC TONE ATHARVAVEDIC INDEPENDENT SVARITA;Mc;0;L;;;;;N;;;;;
1CE2;VEDIC TONE VISARGA SVARITA;Mn;1;NSM;;;;;N;;;;;
1CE3;VEDIC TONE VISARGA UDATTA;Mn;1;NSM;;;;;N;;;;;
1CE4;VEDIC TONE REVERSED VISARGA UDATTA;Mn;1;NSM;;;;;N;;;;;
1CE5;VEDIC TONE VISARGA ANUDATTA;Mn;1;NSM;;;;;N;;;;;
1CE6;VEDIC TONE REVERSED VISARGA ANUDATTA;Mn;1;NSM;;;;;N;;;;;
1CE7;VEDIC TONE VISARGA UDATTA WITH TAIL;Mn;1;NSM;;;;;N;;;;;
1CE8;VEDIC TONE VISARGA ANUDATTA WITH TAIL;Mn;1;NSM;;;;;N;;;;;
1CE9;VEDIC SIGN ANUSVARA ANTARGOMUKHA;Lo;0;L;;;;;N;;;;;
1CEA;VEDIC SIGN ANUSVARA BAHIRGOMUKHA;Lo;0;L;;;;;N;;;;;
1CEB;VEDIC SIGN ANUSVARA VAMAGOMUKHA;Lo;0;L;;;;;N;;;;;
1CEC;VEDIC SIGN ANUSVARA VAMAGOMUKHA WITH TAIL;Lo;0;L;;;;;N;;;;;
1CED;VEDIC SIGN TIRYAK;Mn;220;NSM;;;;;N;;;;;
1CEE;VEDIC SIGN HEXIFORM LONG ANUSVARA;Lo;0;L;;;;;N;;;;;
1CEF;VEDIC SIGN LONG ANUSVARA;Lo;0;L;;;;;N;;;;;
1CF0;VEDIC SIGN RTHANG LONG ANUSVARA;Lo;0;L;;;;;N;;;;;
1CF1;VEDIC SIGN ARDHAVISARGA;Mc;0;L;;;;;N;;;;;
A8E0;COMBINING DEVANAGARI DIGIT ZERO;Mn;230;NSM;;;;;N;;;;;
A8E1;COMBINING DEVANAGARI DIGIT ONE;Mn;230;NSM;;;;;N;;;;;
A8E2;COMBINING DEVANAGARI DIGIT TWO;Mn;230;NSM;;;;;N;;;;;
A8E3;COMBINING DEVANAGARI DIGIT THREE;Mn;230;NSM;;;;;N;;;;;
A8E4;COMBINING DEVANAGARI DIGIT FOUR;Mn;230;NSM;;;;;N;;;;;
A8E5;COMBINING DEVANAGARI DIGIT FIVE;Mn;230;NSM;;;;;N;;;;;
A8E6;COMBINING DEVANAGARI DIGIT SIX;Mn;230;NSM;;;;;N;;;;;
A8E7;COMBINING DEVANAGARI DIGIT SEVEN;Mn;230;NSM;;;;;N;;;;;
A8E8;COMBINING DEVANAGARI DIGIT EIGHT;Mn;230;NSM;;;;;N;;;;;
A8E9;COMBINING DEVANAGARI DIGIT NINE;Mn;230;NSM;;;;;N;;;;;
A8EA;COMBINING DEVANAGARI LETTER A;Mn;230;NSM;;;;;N;;;;;
A8EB;COMBINING DEVANAGARI LETTER U;Mn;230;NSM;;;;;N;;;;;
A8EC;COMBINING DEVANAGARI LETTER KA;Mn;230;NSM;;;;;N;;;;;
A8ED;COMBINING DEVANAGARI LETTER NA;Mn;230;NSM;;;;;N;;;;;
A8EE;COMBINING DEVANAGARI LETTER PA;Mn;230;NSM;;;;;N;;;;;
A8EF;COMBINING DEVANAGARI LETTER RA;Mn;230;NSM;;;;;N;;;;;
A8F0;COMBINING DEVANAGARI LETTER VI;Mn;230;NSM;;;;;N;;;;;
A8F1;COMBINING DEVANAGARI SIGN AVAGRAHA;Mn;230;NSM;;;;;N;;;;;
A8F2;DEVANAGARI SIGN SPACING CANDRABINDU;Lo;0;L;;;;;N;;;;;
A8F3;DEVANAGARI SIGN CANDRABINDU VIRAMA;Lo;0;L;;;;;N;;;;;
A8F4;DEVANAGARI SIGN DOUBLE CANDRABINDU VIRAMA;Lo;0;L;;;;;N;;;;;
```

A8F5;DEVANAGARI SIGN CANDRABINDU TWO;Lo;0;L;;;;N;;;;;
A8F6;DEVANAGARI SIGN CANDRABINDU THREE;Lo;0;L;;;;N;;;;;
A8F7;DEVANAGARI SIGN CANDRABINDU AVAGRAHA;Lo;0;L;;;;N;;;;;
A8F8;DEVANAGARI SIGN ANUSVARA UBHAYATO MUKHA;Lo;0;L;;;;N;;;;;

6. Acknowledgements. This project was made possible in part by a grant from the U.S. National Endowment for the Humanities, which funded the Universal Scripts Project (part of the Script Encoding Initiative at UC Berkeley), and in part by the project entitled “International Digital Sanskrit Library Integration” funded by grant no. 0535207 from the U.S. National Science Foundation.

A. Administrative

1. Title

Summary proposal to encode characters for Vedic in the BMP of the UCS

2. Requester's name

Ireland and UC Berkeley Script Encoding Initiative (Universal Scripts Project)

3. Requester type (Member body/Liaison/Individual contribution)

National Body and Liaison contribution.

4. Submission date

2008-03-06

5. Requester's reference (if applicable)

6. Choose one of the following:

6a. This is a complete proposal

Yes.

6b. More information will be provided later

No.

B. Technical – General

1. Choose one of the following:

1a. This proposal is for a new script (set of characters)

Yes.

1b. Proposed name of script

Vedic Extensions, Devanagari Extended.

1c. The proposal is for addition of character(s) to an existing block

Yes

1d. Name of the existing block

Devanagari.

2. Number of characters in proposal

64 (7, 33, 24).

3. Proposed category (A-Contemporary; B.1-Specialized (small collection); B.2-Specialized (large collection); C-Major extinct; D-Attested extinct; E-Minor extinct; F-Archaic Hieroglyphic or Ideographic; G-Obscure or questionable usage symbols)

Category B.1.

4a. Is a repertoire including character names provided?

Yes.

4b. If YES, are the names in accordance with the “character naming guidelines” in Annex L of P&P document?

Yes.

4c. Are the character shapes attached in a legible form suitable for review?

Yes.

5a. Who will provide the appropriate computerized font (ordered preference: True Type, or PostScript format) for publishing the standard?

Michael Everson.

5b. If available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools used:

Michael Everson, Fontographer.

6a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?

Yes.

6b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?

Yes.

7. Does the proposal address other aspects of character data processing (if applicable) such as input, presentation, sorting, searching, indexing, transliteration etc. (if yes please enclose information)?

Yes.

8. Submitters are invited to provide any additional information about Properties of the proposed Character(s) or Script that will assist in correct understanding of and correct linguistic processing of the proposed character(s) or script. Examples of such properties are: Casing information, Numeric information, Currency information, Display behaviour information such as line breaks, widths etc., Combining behaviour, Spacing behaviour, Directional behaviour, Default Collation behaviour, relevance in Mark Up contexts, Compatibility equivalence and other Unicode normalization related information. See the Unicode standard at <http://www.unicode.org> for such information on other scripts. Also see Unicode Character Database <http://www.unicode.org/Public/UNIDATA/UnicodeCharacterDatabase.html> and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

See above.

C. Technical – Justification

1. Has this proposal for addition of character(s) been submitted before? If YES, explain.

Yes, in a number of documents since the middle of 2006.

2a. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)?

Yes.

2b. If YES, with whom?

Michel Angot, R. Chandrashekar, Malcolm Hyman, Susan Rosenfield, B. V. Venkatakrishna Sastry, Michael Witzel.

2c. If YES, available relevant documents

Specialists in Vedic Sanskrit

3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included?

Indologists, Indo-Europeanists, teachers, students.

4a. The context of use for the proposed characters (type of use; common or rare)

Rare; these are mostly used in Sanskrit dictionaries.

4b. Reference

5a. Are the proposed characters in current use by the user community?

Yes.

5b. If YES, where?

Scholarly publications.

6a. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP?

Yes.

6b. If YES, is a rationale provided?

Yes.

6c. If YES, reference

Accordance with the Roadmap. Keep with other Indic characters.

7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?

No.

8a. Can any of the proposed characters be considered a presentation form of an existing character or character sequence?

No.

8b. If YES, is a rationale for its inclusion provided?

8c. If YES, reference

9a. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters?

No.

9b. If YES, is a rationale for its inclusion provided?

9c. If YES, reference

10a. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character?

No.

10b. If YES, is a rationale for its inclusion provided?

10c. If YES, reference

11a. Does the proposal include use of combining characters and/or use of composite sequences (see clauses 4.12 and 4.14 in ISO/IEC 10646-1: 2000)?

Yes.

11b. If YES, is a rationale for such use provided?

No.

11c. If YES, reference

11d. Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?

No.

11e. If YES, reference

12a. Does the proposal contain characters with any special properties such as control function or similar semantics?

No.

12b. If YES, describe in detail (include attachment if necessary)

13a. Does the proposal contain any Ideographic compatibility character(s)?

No.

13b. If YES, is the equivalent corresponding unified ideographic character(s) identified?

	090	091	092	093	094	095	096	097
0	ै 0900	ऐ 0910	ठ 0920	र 0930	ी 0940	ॐ 0950	ऋ 0960	० 0970
1	ँ 0901	ऑ 0911	ड 0921	ॠ 0931	ु 0941	॑ 0951	ॡ 0961	ॢ 0971
2	ं 0902	ओ 0912	ढ 0922	ल 0932	॒ 0942	॒ 0952	॒ 0962	अँ 0972
3	ः 0903	ओ 0913	ण 0923	ळ 0933	ॢ 0943	े 0953	ॢ 0963	छ 0973
4	ऐ 0904	औ 0914	त 0924	ळ 0934	ॢ 0944	े 0954	। 0964	ॣ 0974
5	अ 0905	क 0915	थ 0925	व 0935	ँ 0945	ँ 0955	॥ 0965	
6	आ 0906	ख 0916	द 0926	श 0936	े 0946		० 0966	
7	इ 0907	ग 0917	ध 0927	ष 0937	े 0947		१ 0967	
8	ई 0908	घ 0918	न 0928	स 0938	ै 0948	क्र 0958	२ 0968	
9	उ 0909	ङ 0919	न 0929	ह 0939	ॉ 0949	ख 0959	३ 0969	ज़ 0979
A	ऊ 090A	च 091A	प 092A		ो 094A	ग 095A	४ 096A	ष 097A
B	ऋ 090B	ॡ 091B	फ 092B		ो 094B	ज 095B	५ 096B	ग 097B
C	ॠ 090C	ज 091C	ब 092C	॑ 093C	ौ 094C	ड़ 095C	६ 096C	ज 097C
D	ँ 090D	झ 091D	भ 092D	ऽ 093D	॑ 094D	ढ़ 095D	७ 096D	१ 097D
E	ऐ 090E	ज 091E	म 092E	ा 093E	ा 094E	फ़ 095E	८ 096E	ड 097E
F	ए 090F	ट 091F	य 092F	ि 093F		य 095F	९ 096F	ब 097F

Based on ISCII 1988**Various signs**

0900	ॐ	DEVANAGARI SIGN INVERTED CANDRABINDU = vaidika adhomukha candrabindu
0901	ँ	DEVANAGARI SIGN CANDRABINDU = anunasika → 0310 combining candrabindu
0902	ं	DEVANAGARI SIGN ANUSVARA = bindu
0903	ः	DEVANAGARI SIGN VISARGA

Independent vowels

0904	अ	DEVANAGARI LETTER SHORT A
0905	आ	DEVANAGARI LETTER A
0906	इ	DEVANAGARI LETTER AA
0907	ई	DEVANAGARI LETTER I
0908	उ	DEVANAGARI LETTER II
0909	ऊ	DEVANAGARI LETTER U
090A	ऋ	DEVANAGARI LETTER VOCALIC R
090C	ॠ	DEVANAGARI LETTER VOCALIC L
090D	ए	DEVANAGARI LETTER CANDRA E
090E	ऐ	DEVANAGARI LETTER SHORT E • for transcribing Dravidian short e
090F	ऑ	DEVANAGARI LETTER E
0910	औ	DEVANAGARI LETTER AI
0911	ओ	DEVANAGARI LETTER CANDRA O
0912	औ	DEVANAGARI LETTER SHORT O • for transcribing Dravidian short o
0913	ओ	DEVANAGARI LETTER O
0914	औ	DEVANAGARI LETTER AU

Consonants

0915	क	DEVANAGARI LETTER KA
0916	ख	DEVANAGARI LETTER KHA
0917	ग	DEVANAGARI LETTER GA
0918	घ	DEVANAGARI LETTER GH
0919	ङ	DEVANAGARI LETTER NG
091A	च	DEVANAGARI LETTER CA
091B	छ	DEVANAGARI LETTER CHA
091C	ज	DEVANAGARI LETTER JA
091D	झ	DEVANAGARI LETTER JHA
091E	ञ	DEVANAGARI LETTER NYA
091F	ट	DEVANAGARI LETTER TTA
0920	ठ	DEVANAGARI LETTER TTHA
0921	ड	DEVANAGARI LETTER DDA
0922	ढ	DEVANAGARI LETTER DDHA
0923	ण	DEVANAGARI LETTER NNA
0924	त	DEVANAGARI LETTER TA
0925	थ	DEVANAGARI LETTER THA
0926	द	DEVANAGARI LETTER DA
0927	ध	DEVANAGARI LETTER DHA
0928	न	DEVANAGARI LETTER NA
0929	न्	DEVANAGARI LETTER NNNA • for transcribing Dravidian alveolar n ≡ 0928 न 093C ण
092A	प	DEVANAGARI LETTER PA
092B	फ	DEVANAGARI LETTER PHA
092C	ब	DEVANAGARI LETTER BA

092D	भ	DEVANAGARI LETTER BHA
092E	म	DEVANAGARI LETTER MA
092F	य	DEVANAGARI LETTER YA
0930	र	DEVANAGARI LETTER RA
0931	ॠ	DEVANAGARI LETTER RRA • for transcribing Dravidian alveolar r • half form is represented as "Eyelash RA" ≡ 0930 र 093C ण
0932	ल	DEVANAGARI LETTER LA
0933	ळ	DEVANAGARI LETTER LLA
0934	ॢ	DEVANAGARI LETTER LLLA • for transcribing Dravidian l ≡ 0933 ळ 093C ण
0935	व	DEVANAGARI LETTER VA
0936	श	DEVANAGARI LETTER SHA
0937	ष	DEVANAGARI LETTER SSA
0938	स	DEVANAGARI LETTER SA
0939	ह	DEVANAGARI LETTER HA

Various signs

093C	्	DEVANAGARI SIGN NUKTA • for extending the alphabet to new letters
093D	ऽ	DEVANAGARI SIGN AVAGRAHA

Dependent vowel signs

093E	ा	DEVANAGARI VOWEL SIGN AA
093F	ि	DEVANAGARI VOWEL SIGN I • stands to the left of the consonant
0940	ी	DEVANAGARI VOWEL SIGN II
0941	ु	DEVANAGARI VOWEL SIGN U
0942	ू	DEVANAGARI VOWEL SIGN UU
0943	ृ	DEVANAGARI VOWEL SIGN VOCALIC R
0944	ॠ	DEVANAGARI VOWEL SIGN VOCALIC RR
0945	ॡ	DEVANAGARI VOWEL SIGN CANDRA E = candra
0946	े	DEVANAGARI VOWEL SIGN SHORT E • for transcribing Dravidian vowels
0947	ै	DEVANAGARI VOWEL SIGN E
0948	॒	DEVANAGARI VOWEL SIGN AI
0949	ॉ	DEVANAGARI VOWEL SIGN CANDRA O
094A	ौ	DEVANAGARI VOWEL SIGN SHORT O • for transcribing Dravidian vowels
094B	ो	DEVANAGARI VOWEL SIGN O
094C	॒	DEVANAGARI VOWEL SIGN AU

Various signs

094D	॑	DEVANAGARI SIGN VIRAMA = halant (the preferred Hindi name) • suppresses inherent vowel
094E	॒	DEVANAGARI VOWEL SIGN PRISHTHAMATRA E • character has historic use only • combines with E to form AI, with AA to form O, and with O to form AU
094F	॒	<reserved>
0950	॑	DEVANAGARI OM
0951	॑	DEVANAGARI STRESS SIGN UDATTA • mostly used for Rigvedic svarita, with rare use for Yajurvedic udatta
0952	॑	DEVANAGARI STRESS SIGN ANUDATTA
0953	॑	DEVANAGARI GRAVE ACCENT

- 0954 ॆ DEVANAGARI ACUTE ACCENT
 0955 ॆ DEVANAGARI VOWEL SIGN CANDRA LONG E
 • used in transliteration of Avestan

Additional consonants

- 0958 क DEVANAGARI LETTER QA
 ≡ 0915 क 093C ॆ
 0959 ख DEVANAGARI LETTER KHHA
 ≡ 0916 ख 093C ॆ
 095A ग DEVANAGARI LETTER GHHA
 ≡ 0917 ग 093C ॆ
 095B ज DEVANAGARI LETTER ZA
 ≡ 091C ज 093C ॆ
 095C ङ DEVANAGARI LETTER DDDHA
 ≡ 0921 ङ 093C ॆ
 095D ढ DEVANAGARI LETTER RHA
 ≡ 0922 ढ 093C ॆ
 095E फ DEVANAGARI LETTER FA
 ≡ 092B फ 093C ॆ
 095F य DEVANAGARI LETTER YYA
 ≡ 092F य 093C ॆ

Additional vowels for Sanskrit

- 0960 ऋ DEVANAGARI LETTER VOCALIC RR
 0961 ॠ DEVANAGARI LETTER VOCALIC LL
 0962 ॡ DEVANAGARI VOWEL SIGN VOCALIC L
 0963 ॢ DEVANAGARI VOWEL SIGN VOCALIC LL

Generic punctuation for scripts of India

These punctuation marks are for common use for the scripts of India despite being named "DEVANAGARI".

- 0964 । DEVANAGARI DANDA
 = purna viram
 • phrase separator
 0965 ॥ DEVANAGARI DOUBLE DANDA
 = deergh viram

Digits

- 0966 ० DEVANAGARI DIGIT ZERO
 0967 १ DEVANAGARI DIGIT ONE
 0968 २ DEVANAGARI DIGIT TWO
 0969 ३ DEVANAGARI DIGIT THREE
 096A ४ DEVANAGARI DIGIT FOUR
 096B ५ DEVANAGARI DIGIT FIVE
 096C ६ DEVANAGARI DIGIT SIX
 096D ७ DEVANAGARI DIGIT SEVEN
 096E ८ DEVANAGARI DIGIT EIGHT
 096F ९ DEVANAGARI DIGIT NINE

Devanagari-specific additions

- 0970 ॰ DEVANAGARI ABBREVIATION SIGN
 0971 ॱ DEVANAGARI SIGN HIGH SPACING DOT

Additional vowel for Marathi

- 0972 अँ DEVANAGARI LETTER CANDRA A
 • Marathi

Various signs

- 0973 ॲ DEVANAGARI SIGN PUSHPIKA
 = vaidika pushpikaa
 • used as a placeholder or “filler”
 • often flanked by double dandas
 0974 ॳ DEVANAGARI CARET
 = vaidika trutikaa
 • zero-advance character centred on the point between two orthographic syllables

Additional consonants

- 0979 ॲ DEVANAGARI LETTER ZHA
 • used in transliteration of Avestan
 097A ॳ DEVANAGARI LETTER HEAVY YA
 • used for an affricated glide
 → 09AF ॳ bengali letter ya

Sindhi implosives

These are added from Amendment 3 to 10646:2003.

- 097B ॳ DEVANAGARI LETTER GGA
 097C ॳ DEVANAGARI LETTER JJA

















































Glottal stop

- 097D ॳ DEVANAGARI LETTER GLOTTAL STOP
 • used for writing Limbu in Devanagari
 • a glyph variant has the connecting top bar





Sindhi implosives

These are added from Amendment 3 to 10646:2003.










- 097E ॳ DEVANAGARI LETTER DDDA
 097F ॳ DEVANAGARI LETTER BBA

	1CD	1CE	1CF
0	 1CD0	 1CE0	 1CF0
1	 1CD1	 1CE1	 1CF1
2	 1CD2	 1CE2	
3	 1CD3	 1CE3	
4		 1CE4	
5	 1CD5	 1CE5	
6	 1CD6	 1CE6	
7	 1CD7	 1CE7	
8	 1CD8	 1CE8	
9	 1CD9	 1CE9	
A	 1CDA	 1CEA	
B	 1CDB	 1CEB	
C	 1CDC	 1CEC	
D	 1CDD	 1CED	
E	 1CDE	 1CEE	
F	 1CDF	 1CEF	



Tone marks for the Samaveda

- 1CD0  VEDIC TONE KARSHANA
= vaidika saamasvara karshanna
- 1CD1  VEDIC TONE SHARA
= vaidika svarita uurdhva shara
- 1CD2  VEDIC TONE PRENKHA
• indicates vibrato
= vaidika saamasvara prenkha
- 1CD3  VEDIC SIGN NIHSHVASA
= vaidika saamasvara yogakaala

Signs for Yajurvedic

- 1CD5  VEDIC TONE YAJURVEDIC AGGRAVATED INDEPENDENT SVARITA
= vaidika svarita adho nyubja
- 1CD6  VEDIC TONE YAJURVEDIC INDEPENDENT SVARITA
= vaidika svarita adhah konna
- 1CD7  VEDIC TONE YAJURVEDIC KATHAKA INDEPENDENT SVARITA
= vaidika svarita adho vakra rekhaa
- 1CD8  VEDIC TONE CANDRA BELOW
= vaidika svarita adhah ardha vakra
- 1CD9  VEDIC TONE YAJURVEDIC KATHAKA INDEPENDENT SVARITA SCHROEDER
= vaidika svarita adho samyukta rekhaa
- 1CDA  VEDIC TONE DOUBLE SVARITA
= vaidika svarita uurdhva dvi rekhaa
- 1CDB  VEDIC TONE TRIPLE SVARITA
= vaidika svarita uurdhva tri rekhaa
- 1CDC  VEDIC TONE KATHAKA ANUDATTA
= vaidika svarita adho rekhaa
- 1CDD  VEDIC TONE DOT BELOW
= vaidika svarita adho bindu

Tone marks for the Satapathabrahmana

- 1CDE  VEDIC TONE TWO DOTS BELOW
= vaidika svarita adho dvi bindu
- 1CDF  VEDIC TONE THREE DOTS BELOW
= vaidika svarita adhas tri bindu





Tone mark for the Rigveda




- 1CE0  VEDIC TONE RIGVEDIC KASHMIRI INDEPENDENT SVARITA
= vaidika uurdhva vakra rekhaa

Tone mark for the Atharvaveda









- 1CE1  VEDIC TONE ATHARVAVEDIC INDEPENDENT SVARITA
= vaidika svarita dvi vakra khanda

Diacritics for visarga

- 1CE2  VEDIC SIGN VISARGA SVARITA
= vaidika madhyerekhaa
- 1CE3  VEDIC SIGN VISARGA UDATTA
= vaidika visarga dakshinnatah uurdhvaga
- 1CE4  VEDIC SIGN REVERSED VISARGA UDATTA
= vaidika visarga vaamatah uurdhvaga
- 1CE5  VEDIC SIGN VISARGA ANUDATTA
= vaidika visarga vaamatah adhoga

- 1CE6  VEDIC SIGN REVERSED VISARGA ANUDATTA
= vaidika visarga dakshinnatah adhoga
- 1CE7  VEDIC SIGN VISARGA UDATTA WITH TAIL
= vaidika visarga dakshinnatah uurdhva vakra
- 1CE8  VEDIC SIGN VISARGA ANUDATTA WITH TAIL
= vaidika visarga vaamatah adho vakra

Marks of nasalization

- 1CE9  VEDIC SIGN ANUSVARA ANTARGOMUKHA
= vaidika anusvaara antarmukha
- 1CEA  VEDIC SIGN ANUSVARA BAHIRGOMUKHA
= vaidika anusvaara naagaphanna
- 1CEB  VEDIC SIGN ANUSVARA VAMAGOMUKHA
= vaidika anusvaara vaamamukha
- 1CEC  VEDIC SIGN ANUSVARA VAMAGOMUKHA WITH TAIL
= vaidika anusvaara vaamamukha savakra
- 1CED  VEDIC SIGN TIRYAK
= vaidika tiryak
- 1CEE  VEDIC SIGN HEXIFORM LONG ANUSVARA
= vaidika anusvaara shattsadrisha
- 1CEF  VEDIC SIGN LONG ANUSVARA
= vaidika anusvaara dakshinnamukha
- 1CF0  VEDIC SIGN RTHANG LONG ANUSVARA
= vaidika anusvaara tthasadrisha

Ardhavisarga

- 1CF1  VEDIC SIGN ARDHAVISARGA
= vaidika jihvaamuliya upadhamaaniya

	A8E	A8F
0	० A8E0	ॐ A8F0
1	१ A8E1	ॡ A8F1
2	२ A8E2	ॢ A8F2
3	३ A8E3	ॣ A8F3
4	४ A8E4	। A8F4
5	५ A8E5	॥ A8F5
6	६ A8E6	० A8F6
7	७ A8E7	ॠ A8F7
8	८ A8E8	ॡ A8F8
9	९ A8E9	
A	अ A8EA	
B	उ A8EB	
C	क A8EC	
D	न A8ED	
E	प A8EE	
F	र A8EF	

Cantillation marks for the Samaveda

A8E0	ॐ	COMBINING DEVANAGARI DIGIT ZERO = vaidika saamasvara anka shuunya
A8E1	ॡ	COMBINING DEVANAGARI DIGIT ONE = vaidika saamasvara anka eka udatta
A8E2	ॢ	COMBINING DEVANAGARI DIGIT TWO = vaidika saamasvara anka dvi svarita
A8E3	ॣ	COMBINING DEVANAGARI DIGIT THREE = vaidika saamasvara anka tri anudatta
A8E4	।	COMBINING DEVANAGARI DIGIT FOUR = vaidika saamasvara anka chatur
A8E5	॥	COMBINING DEVANAGARI DIGIT FIVE = vaidika saamasvara anka pancha
A8E6	०	COMBINING DEVANAGARI DIGIT SIX = vaidika saamasvara anka shatt
A8E7	ॠ	COMBINING DEVANAGARI DIGIT SEVEN = vaidika saamasvara anka sapta
A8E8	ॡ	COMBINING DEVANAGARI DIGIT EIGHT = vaidika saamasvara anka ashta
A8E9	ॢ	COMBINING DEVANAGARI DIGIT NINE = vaidika saamasvara anka nava
A8EA	अ	COMBINING DEVANAGARI LETTER A = vaidika saamasvara abhinihita
A8EB	उ	COMBINING DEVANAGARI LETTER U = vaidika saamasvara u
A8EC	क	COMBINING DEVANAGARI LETTER KA = vaidika saamasvara ka
A8ED	न	COMBINING DEVANAGARI LETTER NA = vaidika saamasvara namana
A8EE	प	COMBINING DEVANAGARI LETTER PA = vaidika saamasvara prannatam
A8EF	र	COMBINING DEVANAGARI LETTER RA = vaidika saamasvara ra
A8F0	ॠ	COMBINING DEVANAGARI LETTER VI = vaidika saamasvara vinata
A8F1	ॡ	COMBINING DEVANAGARI SIGN AVAGRAHA = vaidika saamasvara diirghiiibhaava

Marks of nasalization

A8F2	ॢ	DEVANAGARI SIGN SPACING CANDRABINDU = vaidika candrabindu
A8F3	ॣ	DEVANAGARI SIGN CANDRABINDU VIRAMA = vaidika anusvaara candrabindu tiryak
A8F4	।	DEVANAGARI SIGN DOUBLE CANDRABINDU VIRAMA = vaidika anusvaara dvi candrabindu tiryak
A8F5	॥	DEVANAGARI SIGN CANDRABINDU TWO = vaidika anusvaara candrabindu sadvi
A8F6	०	DEVANAGARI SIGN CANDRABINDU THREE = vaidika anusvaara candrabindu satri
A8F7	ॠ	DEVANAGARI SIGN CANDRABINDU AVAGRAHA = vaidika anusvaara candrabindu saavagraha
A8F8	ॡ	DEVANAGARI SIGN ANUSVARA UBHAYATO MUKHA